Journal of Experimental Botany, Vol. 63, No. 10, pp. 3885–3898, 2012
doi:10.1093/jxb/ers081 Advance Access publication 21 March, 2012
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

T-DNA insertion mutants reveal complex expression patterns of the *aldehyde dehydrogenase 3H1* locus in *Arabidopsis thaliana*

Tagnon D. Missihoun, Hans-Hubert Kirch and Dorothea Bartels*

Institute of Molecular Physiology and Biotechnology of Plants (IMBIO), University of Bonn, Kirschallee 1, 53115 Bonn, Germany

* To whom correspondence should be addressed. E-mail: dbartels@uni-bonn.de

Received 13 December 2011; Revised 17 February 2012; Accepted 20 February 2012

Abstract

The *Arabidopsis thaliana aldehyde dehydrogenase 3H1* gene (*ALDH3H1*; AT1G44170) belongs to family 3 of the plant aldehyde dehydrogenase superfamily. The full-length transcript of the corresponding gene comprises an open reading frame of 1583 bp and encodes a protein of 484 amino acid residues. Gene expression studies have shown that this transcript accumulates mainly in the roots of 4-week-old plants following abscisic acid, dehydration, and NaCl treatments. The current study provided experimental data that the *ALDH3H1* locus generates at least five alternative transcript variants in addition to the previously described *ALDH3H1* mRNA. The alternative transcripts accumulated in wild-type plants at a low level but were upregulated in a mutant that carried a T-DNA insertion in the first exon of the gene. Expression of the transcript isoforms involved alternative gene splicing combined with an alternative promoter. The transcript isoforms were differentially expressed in the roots and shoots and showed developmental stage- and tissue-specific expression patterns. These data support the hypothesis that alternative isoforms produced by gene splicing or alternative promoters regulate the abundance of the constitutively spliced and functional variants.

Key words: abiotic stress, alternative first exon, alternative promoter, alternative splicing, nonsense-mediated mRNA decay.

Introduction

The *Arabidopsis thaliana aldehyde dehydrogenase 3H1* (AT1G44170) gene (*ALDH3H1*) belongs to the aldehyde dehydrogenase (ALDH) superfamily. ALDHs are $NAD(P)^+$-dependent enzymes (ALDH, EC 1.2.1.3), which are widely distributed in all organisms including human and plant genomes (Sophos and Vasiliou, 2003; Stiti *et al.*, 2011). The *Arabidopsis* genome contains 14 members of the *ALDH* gene family encoding proteins belonging to nine families (Kirch *et al.*, 2004). ALDH3H1 belongs to family 3, which has two other members, ALDH3I1 and ALDH3F1. These three proteins share 60% amino acid homology. Phylogenetic analyses have shown that ALDH3H1, ALDH3I1, and ALDH3F1 are separated from other plant ALDHs and betaine aldehyde dehydrogenases (Kirch *et al.*, 2004). ALDH3H1, ALDH3I1, and ALDH3F1 from *A. thaliana* and the family 3 ALDH from the desiccation-tolerant plant species *Craterostigma plantagineum* (Cp-ALDH) were found to be closely related to bacteria and animal ALDHs (Stiti *et al.*, 2011). One feature of the *ALDH3H1* locus is the presence of an intron larger than 1 kb. This distinguishes *ALDH3H1* from all members of the *Arabidopsis ALDH* gene family. Whether this difference has any influence on the expression and function of the gene is so far unclear. The stress-related *Cp-ALDH* gene was found to be induced following abscisic acid (ABA) and dehydration treatments (Kirch *et al.*, 2001). Similar to *Cp-ALDH*, *ALDH3I1* is upregulated following ABA, NaCl, heavy metals, and dehydration treatments, and its ectopic expression in *Arabidopsis* confers abiotic stress tolerance to transgenic plants (Sunkar *et al.*, 2003). Although

Abbreviations: ABA, abscisic acid; ALDH, aldehyde dehydrogenase; EST, expressed sequence tag; FAST, fast *Agrobacterium*-mediated seedling transformation; GUS, β-*glucuronidase*; MDA, malondialdehyde; MS, Murashige and Skoog; SR, serine/arginine-rich; UTR, untranslated region; WT, wild type.

ALDH3F1 is constitutively expressed, plants overexpressing ALDH3F1 performed slightly better under stress conditions than wild-type (WT) plants (Stiti *et al.*, 2011). Gene expression studies have shown that *ALDH3H1* transcripts accumulate mainly in the roots of 4-week-old plants in response to ABA, dehydration, and NaCl (Kirch *et al.*, 2001; Stiti *et al.*, 2011), but the function of *ALDH3H1* remains yet elusive.

Alternative gene splicing confers more flexibility to the transcriptome and has been studied extensively in animal species (Blencowe, 2006). About 20% of *Arabidopsis* and rice genes are predicted to undergo alternative splicing based on expressed sequence tag (EST)/cDNA analyses (Campbell *et al.*, 2006; Wang and Brendel, 2006). More recently, Filichkin *et al.* (2010) reported that a minimum of 42% of intron-containing genes in *Arabidopsis* are alternatively spliced, as deduced from *in silico* analysis. Reports on genes with alternatively spliced transcripts in plants, and particularly in *Arabidopsis*, are likely to increase in the future. Several plant genes have been demonstrated experimentally to undergo alternative splicing. These genes belong to diverse pathways such as primary metabolism, photosynthesis, flowering, and homeostasis under biotic and abiotic stress conditions (Reddy, 2007; Mastrangelo *et al.*, 2011). A subset of these genes encode transcription factors whereby they participate in the regulation of gene expression (Li *et al.*, 2006). One example is the extensive alternative splicing events common to transcripts encoding members of the serine/arginine-rich (SR) protein family, which are known to function in splice site recognition and spliceosome assembly (Reddy, 2007, and references therein). Comprehensive analyses of alternative splicing of pre-mRNAs of *Arabidopsis* SR genes have indicated that 16 of the 20 SR genes displayed alternative splicing events (Palusa *et al.*, 2007; Tanabe *et al.*, 2007). About 95 transcripts are produced from 15 SR genes, thereby increasing the complexity of the SR gene transcriptome 6-fold (Palusa *et al.*, 2007). Although alternative gene splicing has often been reported in plants as a principal component of gene regulation, it is difficult to address whether all alternatively produced transcripts have a biological function. Gene redundancy may further complicate functional analyses, if mutants of the gene do not have a visible phenotype. Careful analyses of alternatively spliced variants are therefore required to understand the biological relevance of splicing events and associated transcripts.

Transcription of some eukaryotic genes can be driven by multiple promoters (Landry *et al.*, 2003; Chen *et al.*, 2007). Each promoter determines a specific transcription start site with a first exon and, accordingly, generates a different alternative first exon transcript. Typically, an alternative first exon is defined as the first exon of one splice variant of a gene, which is either located downstream of a corresponding alternative first exon or absent from other variants (Luzi *et al.*, 2000; Kimura *et al.*, 2006). Using alternative first exons is a variant of alternative splicing that many eukaryotes employ to generate several different transcripts from a single gene (Kornblihtt, 2005; Chen *et al.*, 2007). Alternative first exons can be produced by alternative promoters, alternative splicing of gene transcripts or through a combination of both mechanisms. Reports on alternative first exons are derived mainly from mammalian genomes, especially mouse and human (Landry *et al.*, 2003). Recently, a genome-scale analysis was performed on rice and *Arabidopsis* (Chen *et al.*, 2007), which found that 5.9 and 5% of the total expressed genes in rice and *Arabidopsis* contained alternative first exons, respectively. About 58% of alternative first exon-containing gene structures are derived from alternative promoters. For instance, the *Arabidopsis* RAD21-like gene *SYN1* is involved in meiosis and was shown to produce two transcripts with different 5′ ends, with one transcript starting within the first intron of the other one (Bai *et al.*, 1999). Likewise, the most upstream intron of a rice MADS box gene, *OsMADS1*, was reported to direct gene expression in flowers, whereas the 5′ upstream promoter region lacking the first intron mediated the expression in both vegetative and reproductive tissues (Jeon *et al.*, 2008). Koo *et al.* (2009) reported the existence of alternative promoters within the rice MAP kinase gene *OsBWMK1* and their differential regulation. Although the mechanism is not well understood, the use of alternative promoters and/or alternative gene splicing appears to be a widespread mechanism to increase transcriptome complexity in all organisms. The generation of alternative transcripts is seen as an additional way of creating regulatory diversity and provides a mechanism to synthesize functionally related proteins acting together to mediate a specific biological response (Morello *et al.*, 2002; Parsley and Hibberd, 2006; Qi *et al.*, 2007).

In the current study, when using T-DNA insertion mutants to investigate the function of the *A. thaliana ALDH3H1* gene, it was observed that one mutant line accumulated five alternatively spliced isoforms of *ALDH3H1* (*ALDH3H1-β*, *ALDH3H1-γ*, *ALDH3H1-δ*, *ALDH3H1-ε* and *ALDH3H1-ζ*). These transcript isoforms were also found in WT plants at a lower level in addition to the major and previously known isoform *ALDH3H1-α*. Both the major and the alternative splicing isoforms showed tissue- and developmental stage-dependent expression patterns and were stress inducible. The study revealed the existence of an alternative promoter within the large intron located between the first and second exon. The integrity of this intron was necessary for the correct WT expression of the alternative isoforms. The observations indicated that *ALDH3H1* expression may be regulated through both the use of the alternative promoter and gene splicing. The biological implications of these observations are discussed.

Materials and methods

Plant material and growth conditions

A. thaliana ecotype Col-0 was used throughout this study. WT and transgenic plants were grown on Murashige and Skoog (MS) agar plates (Murashige and Skoog, 1962) or in soil under white light. The soil was a 3:1 mixture of potting soil (Rolfs Gärtnereinkauf,

Siegburg, Germany) and vermiculite. The intensity of the light was approximately 120–150 μE m^{-2} s^{-1} with a 16 h photoperiod at 22 °C. Tissues of 14-day-old and 6-week-old plants were used for gene expression analyses. Inflorescence, leaf, and root tissues were harvested separately after stress treatment and frozen immediately in liquid nitrogen.

Isolation of homozygous T-DNA insertion mutants

Homozygous T-DNA insertion mutants were isolated by two consecutive PCR assays. The first assay involved the use of two gene-specific primers. In the second assay one gene-specific primer was combined with one T-DNA-specific primer. The T-DNA-specific primers S2, T1, and S3 were used for the lines *3h1-A*, *3h1-B*, and *3h1-C* (see Results), respectively. The choice of the gene-specific primer pair was made according to the predicted insertion site and was such that the region to be amplified encompassed the insertion site. Thus, the primer pairs P1/P5 and P1/S2 were used for the line *3h1-A*, primer pairs P1/P5 and P1/T1 for the line *3h1-B* and primer pairs P1/P5 and P1/S3 for the line *3h1-C*. All primers used in this work are listed in Table 1.

RNA isolation and first-strand cDNA synthesis

Isolation of total RNA and RT-PCR analyses were performed as described by Missihoun *et al.* (2011). Gene-specific primers used to amplify first-strand cDNAs are described in Results. Transcripts of the *A. thaliana Actin2* (*AT3G18780*) gene were used as reference (An *et al.*, 1996) and amplified with the primers Ath_Actin 2_fwdM 5′-GGAATCCACGAGACAACCTATAAC-3′ and Ath_Actin 2_rev 5′-AGGAATCGTTCACAGAAAATGTTTC-3′. The parameters of the standard PCR assay were as follows: 5 min at 94 °C, 30 cycles (30 s at 94 °C, 45 s at 62 °C, and 90 s at 72 °C), and 5 min at 72 °C, followed by holding at 4 °C.

Protein-blot analysis

Protein expression was investigated by protein blotting followed by immunodetection using a polyclonal antiserum raised against the ALDH3H1 recombinant protein. Crude protein extracts were separated by 12% SDS-PAGE and electroblotted from the gel onto a nitrocellulose Protran BA-85 membrane (Whatman, Dassel, Germany) at 100 V for 1 h in pre-chilled protein-blot transfer buffer [25 mM Tris/HCl (pH 8.3), 192 mM glycine, 20% (v/v) methanol] (Towbin *et al.*, 1979). The membrane was stained with Ponceau Red solution [0.2% (w/v) Ponceau S in 3% (w/v) trichloroacetic acid] and then blocked for 1 h at room temperature or overnight at 4 °C. The blocking solution contained 0.1% (v/v) Tween 20 and 4% (w/v) not-fat dried milk powder dissolved in 1× TBS [20 mM Tris/HCl (pH 7.5), 150 mM NaCl]. The membrane was probed with ALDH3H1 antiserum diluted 1:5000. The antiserum was produced by BioGenes (Berlin, Germany) from affinity-purified recombinant ALDH3H1 protein and kindly provided by Dr Andrea Ditzer. The immunodetection assay was performed using an ECL Plus Western Blotting Detection Kit (Amersham, Braunschweig, Germany). Signals were detected under a CCD camera (Intelligent Dark Box II; Fujifilm Corp., Tokyo, Japan).

Table 1. Primers used for genotyping the T-DNA insertion mutants and characterizing the spliced isoforms

***ALDH3H1* gene-specific primers**	**Primer sequence (5′ →3′)**
P1	AGAAGGTTTTTGGATCGGCGGA
P2	ATGTTTTACCAACAGAGAGTAC
P3	CAGCTAAAGAACTGGATGGCTC
P4	CGTTTCGCCGGACTATATCTTGACG
P5	TCAACCAACTAAGTCCATGTTTGA
INT	TCTCTTTATGATCTCGCTTTCCCTC
F7	TGCTTCCGTGATTACTCAGAAG
F6	GAGAAATCCAGATATTCGATGCCA
E9R	TGGAAGGAGAGGGCCAAATATTTC
E3R	GAGCCATCCAGTTCTTTAGCTG
I7	GGCATCAATCTTAAGAAGAGAAAAGAG
I5	AACCACACAATAGCGGGAAATGTC
iR	ACAAGCCAAAGACAAAGTCAC
R1	TAGTACTCTCTGTTGGTAAAAC
T-DNA specific primers	
T1	CTGGGAATGGCGAAATCAAGGCATC
S2	CTGGCAAGTGTAGCGGTCAC
S3[a]	TAGCATCTGAATTTCATAACCAATCTCGATACAC

[a] LB3 (Sessions *et al.*, 2002).

Molecular cloning and generation of transgenic plants

Amplicons corresponding to the complete transcript or part of isoforms were generated by PCR using different primer pairs (Table 1). First-strand cDNA was used as template for the PCR. The amplified cDNA fragments were purified directly from the PCR or from agarose gels after electrophoresis using a NucleoSpin® Extract II kit (Macherey-Nagel, Düren, Germany). Purified fragments were cloned into the plasmid vector pJET1.2 provided in the CloneJET™ PCR Cloning Kit (Fermentas; St Leon-Rot, Germany). All fragments were sequenced from recombinant plasmids with the pJET1.2 primers or gene-specific primers. The sequences were aligned with the *ALDH3H1* gene sequence (AT1G44170).

To generate the *ALDH3H1* intron–*β-glucuronidase* (*GUS*) expression cassette, the region of intron 1 (+214 to +1457) between exon 1 and exon 2′ of the *ALDH3H1* gene locus (AT1G44170) was amplified from *A. thaliana* genomic WT (Col-0) DNA using the primers INT-Xba-fwd: 5′-GCGACGTCTAGAGTTAG-GATTCTTTTTTCTCTC-3′ and INT-Xba-rev: 5′-CTGAGCTC-TAGATAAATGAGGAAAGGTCAGTG-3′. The primers were designed with an *Xba*I restriction site (underlined) to facilitate cloning. The expected 1244 bp PCR product was eluted from the agarose gel, digested with *Xba*I, and exchanged with the *ALDH7B4* gene promoter fragment in the plasmid 7gB, a pBin19-derived plasmid, which contains the gene cassette *ALDH7B4-promoter::GUS::nos polyA*. The resulting *3h1-intron::GUS::nos polyA* fusion construct contained the reporter gene *GUS* driven by the *ALDH3H1* intron 1 fragment. The construct was subcloned into *Agrobacterium tumefaciens* cells. Recombinant clones were initially screened by PCR using the primers INT-Xba-rev and pBIN-HindIII: 5′- AGCTATGACCATGATTACGCCAAG-3′. *A. tumefaciens* cells were used to transiently transform *A. thaliana* seedlings by a FAST (fast *Agrobacterium*-mediated seedling transformation) assay (Li *et al.*, 2009).

To generate transgenic plants that ectopically expressed the ALDH3H1 protein, a *Bam*HI cDNA (GenBank accession no. AY072122) fragment (1730 bp) containing the full-length ALDH3H1 coding region was isolated from plasmid pda06974 (RIKEN Institute) by *Bam*HI restriction and subcloned into the *Bam*HI site of the binary pROK2 vector (Baulcombe *et al.*, 1986). The recombinant plasmid was introduced into *A. tumefaciens*. One *A. tumefaciens* clone harbouring the *CaMV35S::ALDH3H1* fusion construct was used to transform *A. thaliana* WT plants (ecotype Col-0) by the floral dip method (Clough and Bent, 1998).

Histochemical detection of GUS activity

In situ detection of GUS activity was performed as described by Jefferson *et al.* (1987). Transiently transformed seedlings were incubated in GUS staining buffer [0.5 mg/ml X-Gluc, 50 mM

NaH_2PO_4 buffer (pH 7.2), 0.1% (v/v) Triton X-100, 8 mM β-mercaptoethanol] at 37 °C for 14–16 h. The tissues were destained in 80% (v/v) ethanol at 80 °C to remove the chlorophyll and kept in 10% (v/v) glycerol. Photographs of the seedlings were taken under a dissecting microscope (SMZ-800; Nikon, Düsseldorf, Germany).

Stress treatments and measurement of malondialdehyde (MDA) content

For gene expression analysis, 15-day-old homozygous seedlings were removed from MS agar plates and either dehydrated on filter paper at room temperature for 1 h or incubated in 200 mM NaCl at room temperature for 4 h. Six-week-old plants were removed from soil and incubated in water or 250 mM NaCl for 6 h. Roots and shoots or leaves were harvested separately and analysed. To assay plant performance under stress, seeds were sown on MS agar containing different concentrations of NaCl. A germination test was performed as described by Jung *et al.* (2008) with minor modifications. Drought stress was imposed on 5-week-old plants by withholding watering for 14 d. Leaf samples were collected for biochemical analyses. To examine the salt tolerance of 5-week-old and soil-grown plants, they were watered every 2 d with 0–300 mM NaCl solutions for 14 d. The degree of peroxidation of membrane lipids was quantified by measuring MDA levels (Missihoun *et al.*, 2011).

Root growth inhibition assay

Approximately 50 seeds were germinated on MS agar containing 2% (w/v) sucrose and grown for 6 d. Seedlings were then transferred to fresh MS agar plates containing 100 mM NaCl. The plates were kept in a vertical position for 3 d. The increase in root length was measured for seedlings of each genotype. Root growth inhibition was estimated as a percentage of the increase in root length on the medium containing NaCl compared with the medium not supplemented with NaCl.

Results

Gene models of ALDH3H1

The *ALDH3H1* (AT1G44170) gene, a member of the aldehyde dehydrogenase family, is composed of ten exons and nine introns. According to the *Arabidopsis* genome resource database TAIR (http://gbrowse.arabidopsis.org/), three protein-coding gene models can be assigned to the gene: AT1G44170.1, AT1G44170.2, and AT1G44170.3. Each gene model would generate a different transcript. The primary transcripts corresponding to the gene models AT1G44170.1 and AT1G44170.2 differ in their 5′ untranslated (5′UTR) sequences (Fig. 1A). An additional intron in the 5′UTR splits the 5′UTR of AT1G44170.1 into two parts, the sequence of which overlaps with the intronless 5′UTR sequence of AT1G44170.2 (Fig. 1A). The protein-coding sequences of AT1G44170.1 and AT1G44170.2 are identical, each initiating at a single translational start codon (ATG) in exon 1, and they will therefore yield the same protein (53 kDa). The transcript model AT1G44170.3 predicts a shorter transcript, which lacks the first exon of the locus and would therefore code for a smaller protein (46 kDa) with a different N terminus. The gene model AT1G44170.3 suggests the existence of a short exon (exon 2′) of 24 bp at the 5′ end including a small 5′UTR in the first intron of the locus (Fig. 1A). Previous analyses did not focus on transcripts associated with each of these gene models and therefore the transcripts corresponding to AT1G44170.1 and AT1G44170.2 were studied as a single *ALDH3H1* transcript encoding a 53 kDa protein (Kirch *et al.*, 2001, 2005). The *Arabidopsis* EST clone DR182505 (Alexandrov *et al.*, 2006) confirms the *in silico* prediction for AT1G44170.3. The use of different T-DNA insertion mutants allowed us to detect and analyse the expression patterns of several *ALDH3H1* transcript isoforms including the one predicted by the gene model AT1G44170.3. For a better coherence of the text, transcripts corresponding to the models AT1G44170.1 and AT1G44170.2 are considered as the major isoform and are hereafter named *ALDH3H1-α*, whereas that corresponding to AT1G44170.3 is termed *ALDH3H1-β*.

T-DNA insertions reveal complex expression patterns of the ALDH3H1 *gene*

Two T-DNA insertion mutants (depicted in Fig. 1B) were identified from the Syngenta Arabidopsis Insertion Library collection (SAIL; McElver *et al.*, 2001; Sessions *et al.*, 2002) and seeds were obtained from the European Arabidopsis Stock Centre (NASC). The line SAIL_832_A05, hereafter named *3h1-A*, carried a T-DNA insertion in exon 1 and the line SAIL_828_D05, hereafter named *3h1-C*, had an insertion in exon 5 (red triangles in Fig. 1B). The lines *3h1-A* and *3h1-C* carried the same T-DNA (4763 bp) derived from plasmid pDAP101 (McElver *et al.*, 2001; Sessions *et al.*, 2002). A third *ALDH3H1* T-DNA insertion, the line *KO69*, hereafter renamed *3h1-B*, was obtained from the collection described by Rìos *et al.* (2002) and carried a T-DNA fragment (7355 bp) in the first intron of *ALDH3H1* (yellow triangle in Fig. 1B). The T-DNA fragment in the line *3h1-B* was derived from plasmid pPCV6NFHyg (Mathur *et al.*, 1998). Homozygous plants were isolated for each T-DNA mutant (Fig. 2A); transcripts from the lines were examined by RT-PCR with different primer combinations (for primers, see Fig. 1B and Table 1). The homozygous *3h1-A* plants were initially screened using the gene-specific primers P1 and P5 to amplify the full-length transcript *ALDH3H1-α*. The results showed that line *3h1-A* did not accumulate the expected transcript *ALDH3H1-α*. Putative transcripts among the RT-PCR products were detected as a single band on agarose gels when primer P1 was replaced by primer P4, which is located downstream of the T-DNA insertion site (Figs 1B and 2B). This observation was examined in detail. We hypothesized that this amplicon could correspond to the transcript variant *ALDH3H1-β* based on the gene model predictions on the TAIR webpage. Two additional primers were designed for RT-PCR experiments to identify the transcripts. One primer (P2) covered the first 22 nt of exon 2′ and the other primer (P3) was located in exon 2 (Fig. 1B). Primer P2 was designed to amplify specifically the transcript *ALDH3H1-β*, whereas primers P3 and P4 would not discriminate *ALDH3H1* transcript isoforms with variants upstream of exon 2.

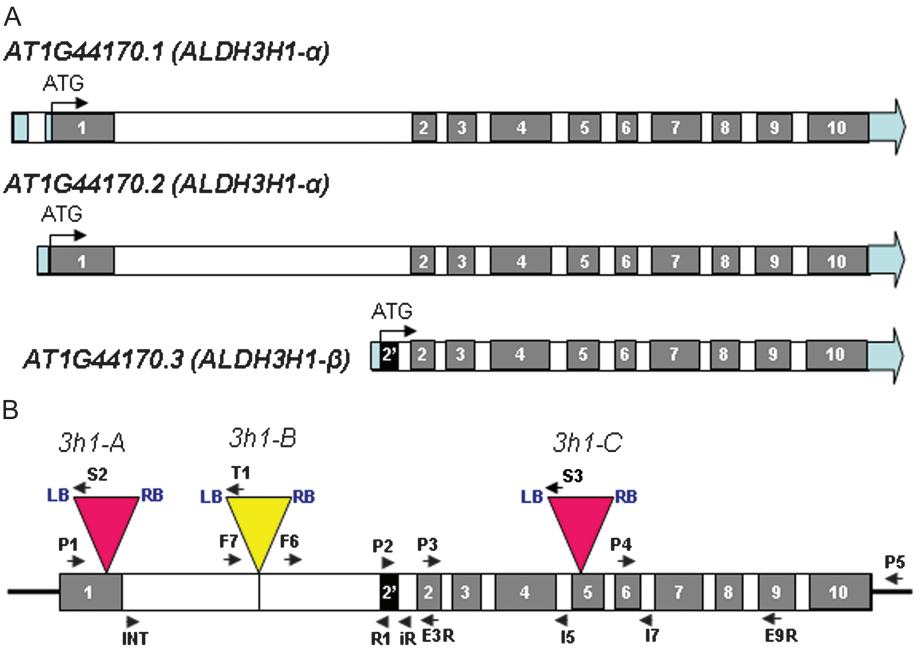

Fig. 1. Structural organization of the *ALDH3H1* gene locus in *A. thaliana*. (A) Schematic representation of the protein-coding gene models of the *ALDH3H1* locus. Grey boxes indicate the protein-coding regions of exons numbered from 1 to 10. The black box represents exon 2′. Open white boxes represent introns. Blue boxes and arrows represent the 5′- and 3′UTR, respectively. The black arrow labelled ATG indicates the position of the translational start codon. (B) Locations of the T-DNA insertions and the primers used for molecular analyses. Inverted red and yellow triangles indicate T-DNAs. The names above the triangles (*3h1-A*, *3h1-B*, and *3h1-C*) refer to the corresponding T-DNA insertion lines. The 5′- and 3′UTRs are shown as horizontal black lines. Arrows indicate the orientation of primers and their approximate position within the gene or the T-DNA. LB and RB indicate the left and right borders of the T-DNA fragment. See Table 1 for primer sequences.

Consistent with our hypothesis, an amplicon of the expected size was obtained from the homozygous *3h1-A* mutant using primers P2 and P5 (Fig. 2C). To verify that the amplicon was also present in the WT, the primer pairs P2/E9R and P2/P5 were used in RT-PCR experiments using total RNA from WT plants with varying PCR cycle numbers. The results showed that the same amplicons were produced in WT and *3h1-A* plants (Fig. 2D).

Another RT-PCR-based approach was used to map the length of the transcript from which the amplicon was derived. Primers F6 and F7 are further upstream of exon 2′ and were included in the transcript analyses. The results revealed the presence of additional amplicons (Fig. 3). Whether these amplicons represented some alternative transcripts generated through alternative splicing or were PCR artefacts was examined. Primers were designed to obtain cDNA fragments spanning the gene from exon 1 to a few nucleotides beyond exon 10 (Fig. 1B; Table 1). RT-PCR experiments were conducted using total RNA from WT and *3h1-A* plants, and the amplicons were cloned and sequenced. The alignment of DNA sequences of cDNA fragments obtained by RT-PCR to the gene sequence confirmed the existence of additional alternative splicing variants of the *ALDH3H1* gene, mostly generated through intron-retention-type alternative splicing events. As shown in Fig. 4, five novel splicing variants of the *ALDH3H1* primary mRNA were identified and named: *ALDH3H1-β*, *ALDH3H1-γ*, *ALDH3H1-δ*, *ALDH3H1-ε* and *ALDH3H1-ζ*. Because clones were chosen randomly during the cloning procedure, it could not be excluded that additional splicing variants occurring at low frequency were not detected.

Only full-length ALDH3H1 protein accumulates in A. thaliana

The nucleotide sequences of the five newly isolated isoforms were analysed *in silico*. Four of them contain at least one premature termination codon that would yield a truncated version of the ALDH3H1 protein. The locations of these premature termination codons are indicated by asterisks in Fig. 4. Only isoform *ALDH3H1-β* corresponding to the gene model AT1G44170.3 did not harbour any premature termination codon and would therefore encode a truncated ALDH3H1 protein (ALDH3H1-*β*). Depending on the translation start codon used, the putative protein ALDH3H1-*β* is predicted to be 421 aa (46.2 kDa) or 395 aa (43 kDa). The accumulation of ALDH3H1-*β* was investigated. Equal amounts of protein extracts from the WT, the homozygous *3h1-A*, *3h1-B*, and *3h1-C* mutants, and two independent lines overexpressing the protein ALDH3H1-*α* were analysed by protein blots followed by immunodetection using antibodies raised against ALDH3H1. In comparison with the WT and the ALDH3H1 overexpressors, no expression of the ALDH3H1 protein (53.2 kDa) was found in the T-DNA insertion mutants (Fig. 5). In addition, no truncated protein of the size of ALDH3H1-*β* was identified. The extra band

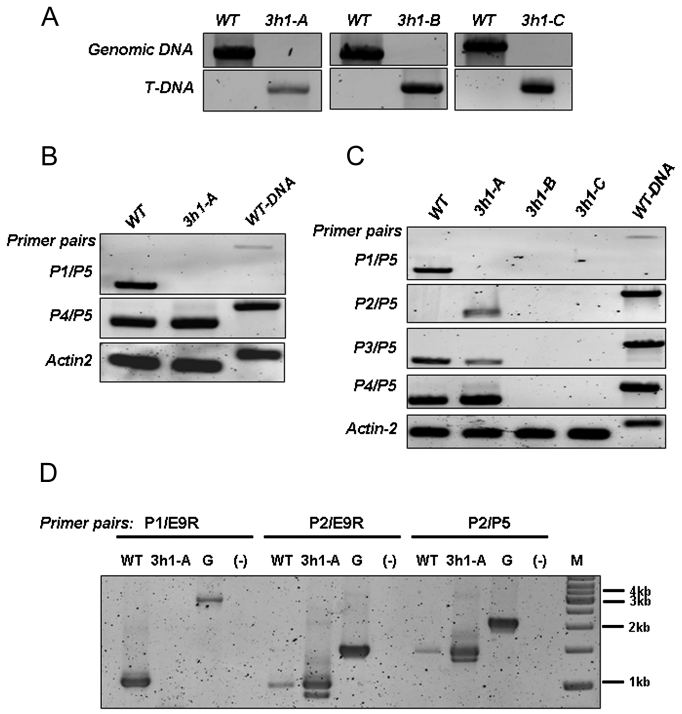

Fig. 2. Molecular analysis of the *ALDH3H1* T-DNA insertion mutants *3h1-A, 3h1-B*, and *3h1-C*. (A) PCR-based genotyping of the homozygous *3h1-A, 3h1-B* and *3h1-C* lines. The upper and lower panels show amplicons produced using *ALDH3H1* gene-specific primers and the combination of gene and T-DNA specific primers in PCR experiments using genomic DNA, respectively. See Materials and methods for details of the genotyping. WT, genomic DNA (Col-0). (B–D) Comparative analysis of *ALDH3H1* transcripts in homozygous *3h1-A, 3h1-B*, and *3h1-C* mutants. P1, P2, P3, and P4 are the forward primers used in combination with the reverse primer P5 or E9R in the RT-PCR assays. Primer pair P1/P5 was specific to *ALDH3H1-α*, whereas P2/P5 and P2/E9R allowed amplification of *ALDH3H1-β*. Primer pairs P3/P5 and P4/P5 could not discriminate between the variants. Genomic WT DNA (Col-0) was used as control (G), while (–) denotes a control reaction without nucleic acid. M, DNA size marker. See Fig. 1B and Table 1 for details of the primers.

marked by an asterisk in Fig. 5 was intrinsic to the antiserum and not to ALDH3H1 because this band occurred in all samples including that from the knockout line *3h1-C*, which does not produce a transcript. These observations indicated that the transcript *ALDH3H1-β* may not be translated into a protein or that it accumulates at a level below detection.

An alternative promoter and alternative splicing regulate expression of the ALDH3H1 *gene at the transcription level*

No amplicon was obtained from the RT-PCR assays when primer F7 was replaced by primer INT, which is located about 400 bp upstream of F7 (Fig. 1B). This suggests that either the 5′ end of the newly identified alternative *ALDH3H1* transcripts locates between these primers or

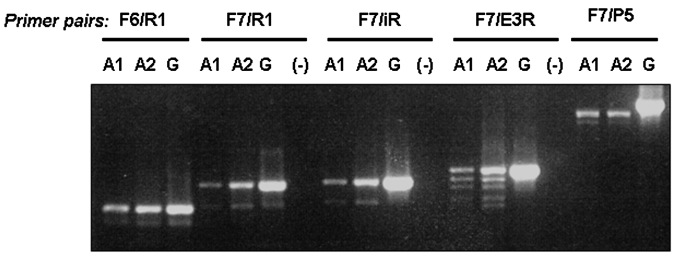

Fig. 3. Agarose gel electrophoresis of amplicons derived from RT-PCR experiments using total RNA from homozygous *3h1-A* plants. A1 and A2 are two independently extracted RNA samples, G denotes genomic WT DNA (Col-0) and (–) a control reaction without nucleic acid. See Fig. 1B for primer locations.

other splicing events occur in the region between the primers INT and F7. In agreement with this, the existence of a 5'-end EST clone (DR182505; Alexandrov *et al.*, 2006) that corresponds to the isoform *ALDH3H1-β* described in this work suggests that isoforms may exist that are produced from an alternative promoter located within intron 1 upstream of exon 2′ (Fig. 1B). To investigate whether a functional alternative promoter is present within the *ALDH3H1* locus, the expression pattern of the major isoform *ALDH3H1-α* was first compared with that of the alternative isoforms. Primer P2 (located in exon 2′; Fig. 1B) was used to discriminate between *ALDH3H1-α* and the alternative variants because exon 2′ is present only in the latter. When homozygous and heterozygous *3h1-A* plants were examined by RT-PCR, the heterozygous plants (*3h1-Aa*) expressed both *ALDH3H1-α* and the alternative transcripts, in contrast to homozygous (*3h1-AA*) plants, which expressed only the alternative transcripts (Fig. 6A). The strength of the signal obtained from *3h1-Aa* samples was about half of that derived from the homozygous *3h1-AA* plants, suggesting that the presence of the T-DNA in exon 1 could have enhanced expression of the alternative transcripts. This is supported by the observation that the WT accumulated lower levels of the alternative transcripts than the homozygous *3h1-A* plants (Fig. 2D). Transcription of the major and alternative transcript variants may be differentially activated; therefore, the relative abundance of the isoforms was studied in a subsequent experiment. Semi-quantitative RT-PCR assays were conducted using total RNA from WT and homozygous *3h1-A* plants. The primer pairs P1/E9R and P2/E9R were used to amplify *ALDH3H1-α* and the alternative isoforms, respectively. As shown in Fig. 6B (reactions R1 and R2), the WT accumulated more of the isoform *ALDH3H1-α* than the alternative isoforms. In contrast, higher levels of the alternative isoforms were present in *3h1-A* mutant plants than in the WT (Fig. 6C).

When different organs were examined, both developmental-dependent and tissue-specific expression patterns were observed among the variants (Fig. 6D). The major isoform *ALDH3H1-α* was present but at a lower level in WT seedlings than in fully expanded leaves and flowers. The WT contained lower amounts of the alternative isoforms in seedlings and leaf tissues than in flowers. In contrast, the

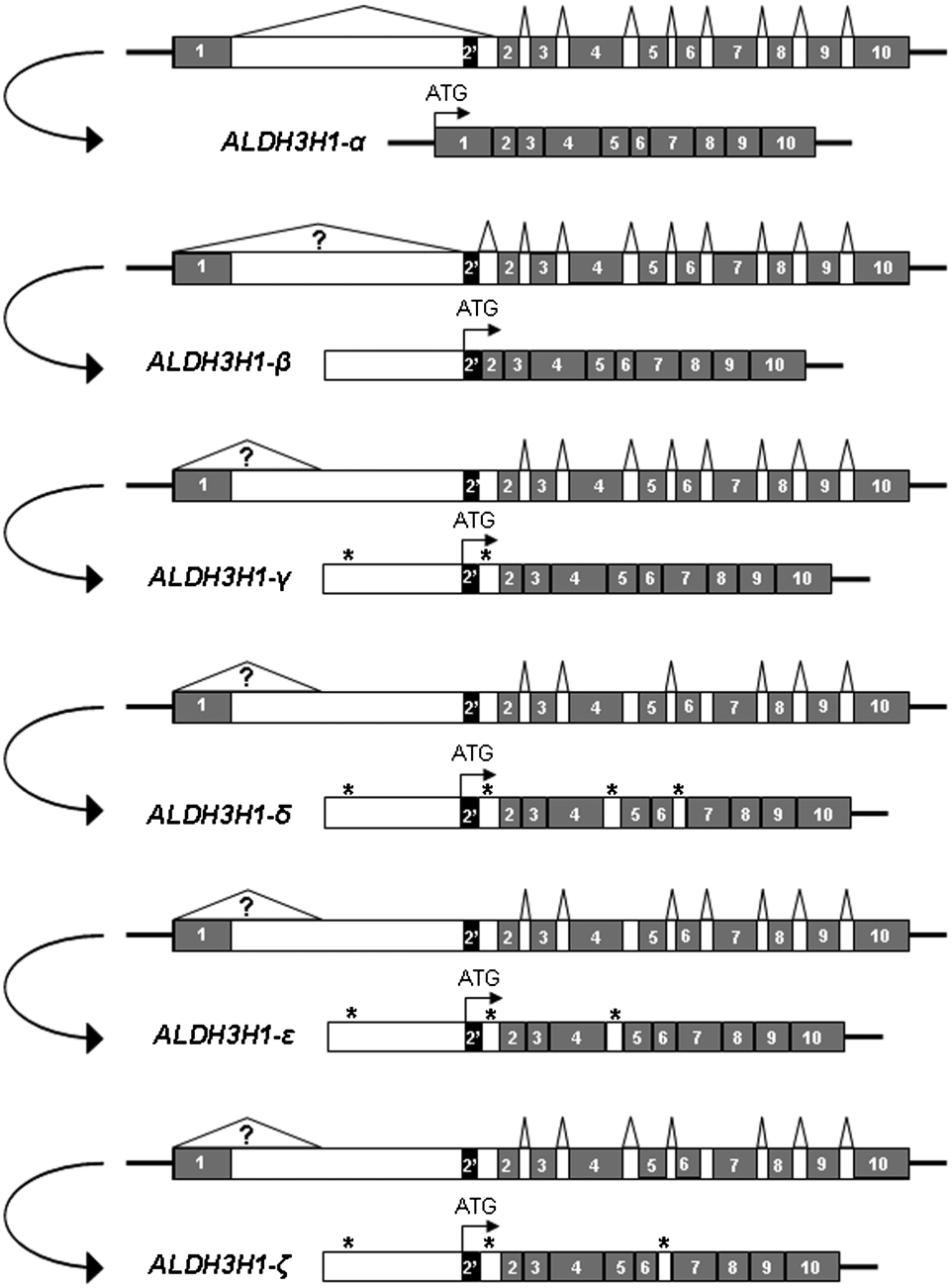

Fig. 4. Schematic representation of the major spliced isoforms of *ALDH3H1*. The structure of the primary transcript and the splicing events leading to the isolated spliced transcripts are shown above each variant. Grey boxes indicate the protein-coding region of exons. The black box represents exon 2′. Open white boxes indicate introns. Horizontal black lines indicate the 5′- and 3'UTR. The 5′UTR is not shown for the alternative isoforms because it could not be determined unequivocally. Further details are provided in the text. Triangles indicate splicing events confirmed by identification of the resulting cDNAs, while question marks indicate a putative splicing event. An asterisk (*) indicates the presence of at least one stop codon in the retained intron. The black arrow labelled ATG indicates the position of the translational start codon of the predicted longest open reading frame.

alternative isoforms were upregulated in all examined tissues of the *3h1-A* line. Differences in expression were also seen among isoforms. For example, amplicon 1 appeared to be downregulated in flowers, whereas amplicon 2 was constitutively expressed in all tissues (Fig. 6D, panel P2/E9R). A similar differential expression was observed for the isoforms δ, ϵ and ζ (Fig. 6D, panels P2/I5 and P2/I7). These observations indicate that isoforms are differentially regulated and support the hypothesis that these transcripts are expressed from different promoters.

To test the activity of a putative alternative promoter, the entire intron region (+214 to +1457) between exon 1 and exon 2' was amplified from WT *A. thaliana* (Col-0) genomic DNA and subcloned in front of the reporter gene *GUS* to generate the expression cassette *3h1-intron::GUS::nospolyA*. This construct was functionally analysed in *A. thaliana*

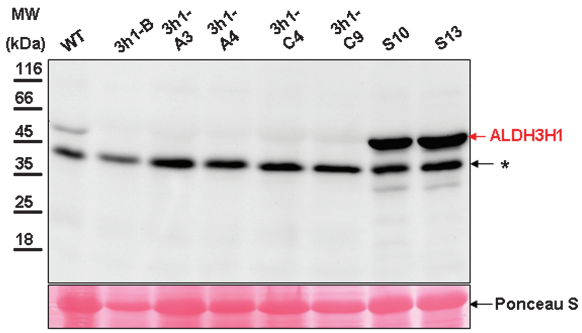

Fig. 5. Accumulation of the ALDH3H1 protein in WT, ALDH3H1 overexpressors, and T-DNA insertion mutants. Fifteen micrograms of crude protein extract from plant leaves was analysed by immunoblotting. Antiserum raised against the ALDH3H1 protein was used for the immunodetection assay. Line 3h1-B is homozygous *3h1-B*, lines 3h1-A3 and 3h1-A4 are homozygous *3h1-A* plants, and lines 3h1-C4 and 3h1-C9 are homozygous *3h1-C* plants. S10 and S13 are two independent transgenic lines overexpressing the protein ALDH3H1-α. The extra band indicated by an asterisk (*) is intrinsic to the antiserum and is not ALDH3H1 specific because this band occurred in all samples including that from the knockout line *3h1-C*, which does not produce a transcript. (This figure is available in colour at *JXB* online.)

seedlings. Transiently transformed seedlings were found to express the reporter protein in roots and cotyledons (Fig. 7). *In silico* analysis of intron 1 of the *ALDH3H1* locus using PLACE Web Signal Scan (Prestridge, 1991; Higo *et al.*, 1999) revealed the presence of several *cis* elements. Two dehydration-responsive element/C-repeat–low temperature-responsive elements (DRE/CRT) and one ACGT box (a putative ABA responsive element) were found at 366, 314, and 333 nt upstream of the predicted translational start codon ATG in exon 2′, respectively. Using the PlantCare tool for *in silico* analysis of promoter sequences (Lescot *et al.*, 2002), the closest predicted TATA box and the corresponding transcription start signal were found at 133 and 103 nt, respectively, upstream of the predicted translational start codon ATG in exon 2'. Only two nucleotides separate the predicted transcription start site from the first nucleotide of the 5′ EST clone DR182505 of *ALDH3H1* (Alexandrov *et al.*, 2006). This indicates that the intronic fragment contains functional *cis* elements able to direct the transcription of a downstream protein-coding DNA fragment. Analysis of the genomic DNA sequence in this locus revealed that the classical plant mRNA splicing acceptor site does not exist at the junction between the intron sequence and exon 2′. The two nucleotides TA terminate the intronic region upstream of the exon 2′ instead of AG found in the consensus sequence (GT…AG) at the intron–exon junction of plant protein-coding genes (Reddy, 2007). Our findings support the existence of an alternative promoter within intron 1. This promoter may direct expression of the *ALDH3H1* alternative transcripts, and particularly that of *ALDH3H1-β*.

Examination of dehydration and salt stress responses in the ALDH3H1 *mutants*

Several *Arabidopsis ALDH* genes have been shown to be stress responsive and their ectopic expression in transgenic plants leads to stress tolerance (Sunkar *et al.*, 2003; Kirch *et al.*, 2005; Kotchoni *et al.*, 2006). Although no protein corresponding to the alternative transcript *ALDH3H1-β* had been identified in WT or mutant plants so far, it is possible that *ALDH3H1-β*, in addition to the other alternative transcript variants, is functional at the RNA level. To test whether the alternative *ALDH3H1* transcript isoforms were stress responsive, the accumulation patterns were examined in *3h1-A* plants in response to dehydration or NaCl treatment. Fifteen-day-old homozygous *3h1-A* seedlings or 6-week-old heterozygous *3h1-A* plants were used to simultaneously monitor *ALDH3H1-α* (the major isoform) and the alternative *ALDH3H1* isoforms. Primers P1 and P2 were used in combination with primer P5 in the RT-PCR assays as described above. Primer pair P1/P5 amplified the major variant, P2/P5 the alternative isoforms, and P4/P5 all isoforms. As shown in Fig. 8A, the alternative isoforms accumulated at a low level in seedling roots following NaCl treatment but decreased in the shoot. In adult plants, they were upregulated by salt treatment in both roots and leaves. In contrast, *ALDH3H1-α* did not increase in the roots but increased in the leaves following salt treatment (Fig. 8B).

Based on these observations, the contribution of the alternative isoforms was investigated further with respect to abiotic stress. The salt stress response was examined in the mutant line *3h1-A*, which accumulates high levels of the alternative isoforms compared with the WT. Line *3h1-C*, which is a null-mutant of *ALDH3H1*, was also included in the experiment. Inhibition of root growth was assayed for all genotypes. As shown in Fig. 9A, homozygous *3h1-A* (*3h1-AA*) and *3h1-C* mutants were affected in their root growth to the same extent, and this was more severe than in the WT and the heterozygous *3h1-Aa*, which showed similar phenotypes. These observations indicated that the alternative transcript variants do not functionally compensate for the lack of the major isoform *ALDH3H1-α* under salt stress, at least for root growth inhibition by salt. The results also showed that mutants lacking *ALDH3H1-α* are sensitive to salt stress.

Because the major transcript *ALDH3H1-α* was also responsive to stress, it was ectopically expressed in stably transformed *A. thaliana* plants. Several independent transgenic lines that overexpress the full-length ALDH3H1 protein were obtained. Results of the protein-blot analysis for two representative lines (S10 and S13) are shown in Fig. 5. The WT and transgenic plants were sown on MS agar medium containing different concentrations of NaCl and the germination rate was scored. The WT and transgenic plants were similarly affected in their germination rate and seedling growth (Fig. 9B). No difference was seen between *3h1-B* and WT plants with regard to seed germination rate on medium supplemented with NaCl.

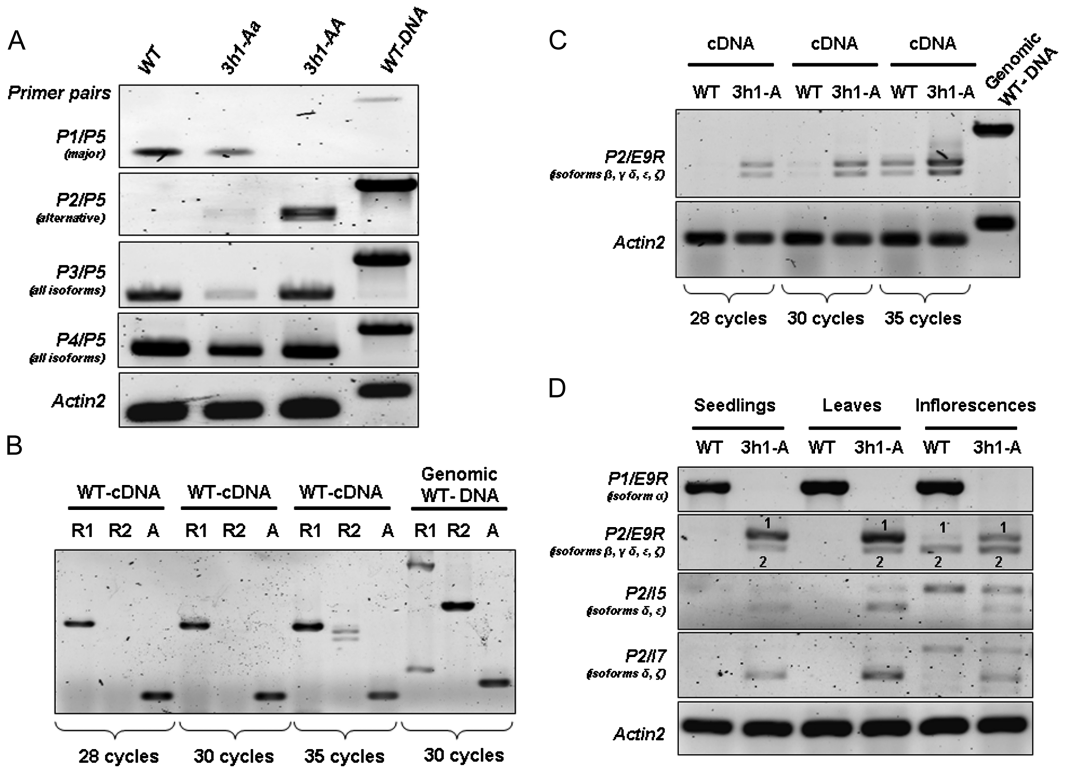

Fig. 6. RT-PCR-based expression analysis of the major and alternatively spliced *ALDH3H1* isoforms under normal growth conditions. Primer pairs that were used in the RT-PCR assays are shown for each panel to refer to the major (*ALDH3H1-*α) and the alternative (β, γ, δ, ε, and ζ) isoforms. Primer pair P1/P5 was specific to *ALDH3H1-*α, whereas P2/P5 and P2/E9R allowed simultaneous amplification of *ALDH3H1-*β and the other alternative isoforms (γ, δ, ε, and ζ). Primer pairs P3/P5 or P4/P5 could not discriminate between the variants (see Fig. 1b for primer locations). (A) Comparative analysis of the accumulation of *ALDH3H1* transcripts in WT, homozygous (*3h1-AA*) and heterozygous (*3h1-Aa*) *3h1-A* mutants. (B, C) Semi-quantitative analysis of the expression of the isoforms in WT cDNA (B) or both WT and *3h1-A* fully developed leaf tissues (C). R1, R2, and A in (B) represent reactions involving primer pairs P1/E9R and P2/E9R and those for *Actin2*, respectively. (D) Tissue-specific expression of the isoforms. Primer pairs P2/I5 and P2/I7 partially discriminated between the *ALDH3H1* isoforms δ, ε, and ζ and did not amplify the isoforms α, β, or γ. The numbers 1 and 2 indicate amplicons 1 and 2, respectively.

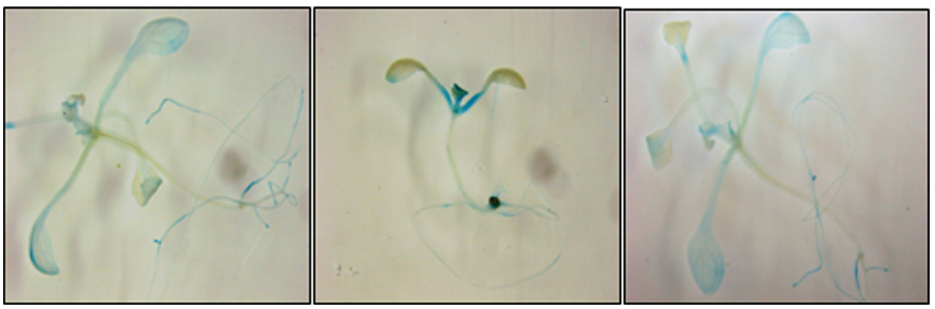

Fig. 7. *In situ* GUS staining of transiently transformed *A. thaliana* seedlings carrying the *3h1-intron::GUS* gene cassette. The promoter activity of the intron 1 fragment of *ALDH3H1* is shown by expression of the GUS reporter protein in roots and cotyledons. The figure shows GUS staining from three different seedlings.

However, line S13 similar to the other tested ALDH3H1-overexpressing lines, accumulated less MDA, a measure of the extent of membrane peroxidation, than the WT and significantly less than the mutant *3h1-B* (Fig. 9C). The salt stress response was also examined in adult WT and ALDH3H1-overexpressing plants. Similar MDA contents were found in all lines after 14 d of salt stress (Fig. 9D). When the plants were subjected to drought stress for 14 d, the ALDH3H1 overexpressors accumulated less MDA than the WT and the *3h1-B* mutant (Fig. 9E). However, the fitness of the overexpressors was not improved, as they all dried out, like the WT. No difference was seen between the WT and transgenic plants when treated with methyl viologen, a reactive oxygen species propagator (data not shown). In summary, overexpression of the ALDH3H1 protein did not confer abiotic stress tolerance to *Arabidopsis*. The results,

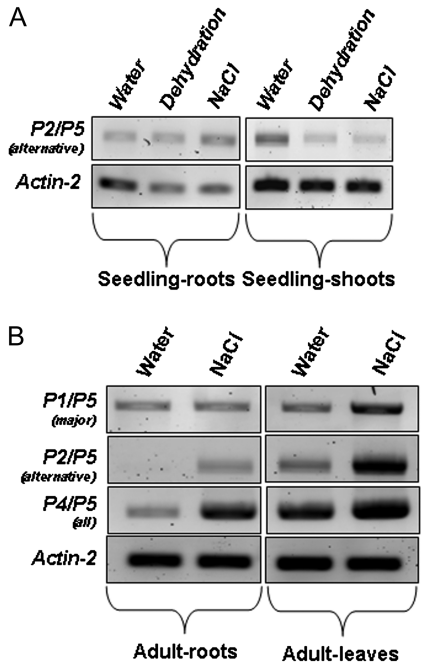

Fig. 8. Expression patterns of the *ALDH3H1* transcript isoforms in response to stress conditions. The expression pattern of the major transcript (*ALDH3H1-α*) was compared with that of the alternatively spliced *ALDH3H1* isoforms (*β*, *γ*, *δ*, *ε*, and *ζ*) following dehydration and NaCl treatments in seedlings (A) and in adult plants (B). Primer pairs that were used in the PCR assays are shown for each panel to refer to the major (*ALDH3H1-α*) and the alternative isoforms. Primer pair P1/P5 was specific to *ALDH3H1-α* whereas P2/P5 allowed simultaneous amplification of *ALDH3H1-β* and the other alternative isoforms. Primer pair P4/P5 could not discriminate between the variants (see Fig. 1B for the primer locations).

however, support a role of the ALDH3H1 protein in alleviating stress injuries related to membrane lipid peroxidation under moderate stress conditions.

Discussion

This study has provided an analysis of the *A. thaliana ALDH3H1* gene utilizing T-DNA insertion mutants. Five alternatively spliced transcript isoforms, namely *ALDH3H1-β*, *-γ*, *-δ*, *-ε* and *-ζ*, were identified, in addition to the major variant, *ALDH3H1-α*. This finding contributes to the increasing number of alternatively spliced genes found in *A. thaliana*. While exon skipping is the most frequent type of alternative splicing in animal cells, intron retention is dominant in plants (Ner-Gaon *et al.*, 2004; Kim *et al.*, 2007). Hence, the risk is high of taking incompletely spliced transcripts (due to splicing errors) as biologically relevant splicing variants. Conversely, alternative transcript isoforms with biological properties can easily be mistaken as transcriptome junk. Examination of alternative gene splicing therefore requires careful experimental approaches. Nearly all intron-containing genes could produce splice variants if most isoforms are generated from random splicing errors (Reddy, 2007). Co-purification of some intron-retaining mRNAs with ribosomes has supported intron retention as a valid splicing mechanism and has underlined the importance of alternative splicing in plant gene regulation (Ner-Gaon *et al.*, 2004). Alternative splicing of some genes is found to be conserved between evolutionarily distant plant species. A comparison of rice and *Arabidopsis* showed that about 27% of alternatively spliced genes have the same alternative splicing type in both species (Wang and Brendel, 2006). However, not all splicing events are conserved; some may have evolved later in evolution and, as such, they are species specific. This is probably the case for *ALDH3H1*, which structurally differs from other *Arabidopsis ALDH* genes. A unique feature of *ALDH3H1* is the 1.2 kb intron that separates exons 1 and 2 (Fig. 1). Alignment of the sequence of intron 1 with sequences deposited in the plant DNA sequence databases did not identify any related sequences, which points to a recent evolutionary event. The position and size of intron 1 are likely to influence expression of the *ALDH3H1* gene. Whether the size of intron 1 has facilitated alternative splicing is unclear.

Among the five alternative *ALDH3H1* isoforms described here, only *ALDH3H1-β* does not derive from intron retention-type alternative splicing (Fig. 4). It is also unlikely that it is generated through skipping of exon 1, because it was detected in both the WT and *3h1-A*, despite the presence of T-DNA in the latter (Fig. 1B). The reading frame of *ALDH3H1-β* is identical to the spliced isoform *ALDH3H1-α* starting downstream of exon 2 (Figs 1A and 4). These features argue for *ALDH3H1-β* being the product of the gene model AT1G44170.3 and being identical to the 5′-end EST clone DR182505 reported by Alexandrov *et al.* (2006). The 5′UTR sequence of *ALDH3H1-β* and that of the other alternative splicing variants have not been examined. Such an analysis would not be informative because of the extensive overlap upstream of exon 2′ between *ALDH3H1-β* and the other isoforms (Fig. 4). The 5′UTR of *ALDH3H1-β* and *ALDH3H1-γ*, *-δ*, *-ε*, and *-ζ* could be partially overlapping.

A protein derived from *ALDH3H1-β* was not detected, although *ALDH3H1-β* does not contain any premature termination codon in the protein-coding region. This can be attributed to a lower stability of *ALDH3H1-β* transcripts undergoing a post-transcriptional degradation that prevents their translation into protein. Another reason could be that the translation efficiency of *ALDH3H1-β* is poor and, accordingly, the concentration of ALDH3H1-*β* would be below the detection limit. Alternatively, the absence of ALDH3H1-*β* protein could indicate that the occurrence of the alternative isoforms has functions other than the diversification of protein variants. The four other isoforms (*ALDH3H1-γ*, *-δ*, *-ε*, and *-ζ*) were found to contain at least one premature termination codon, which makes them putative targets of the nonsense-mediated mRNA decay surveillance mechanism (Belgrader *et al.*, 1994; Maquat, 2004; Palusa and Reddy, 2010). Accumulating reports support the idea that the accumulation of unproductive

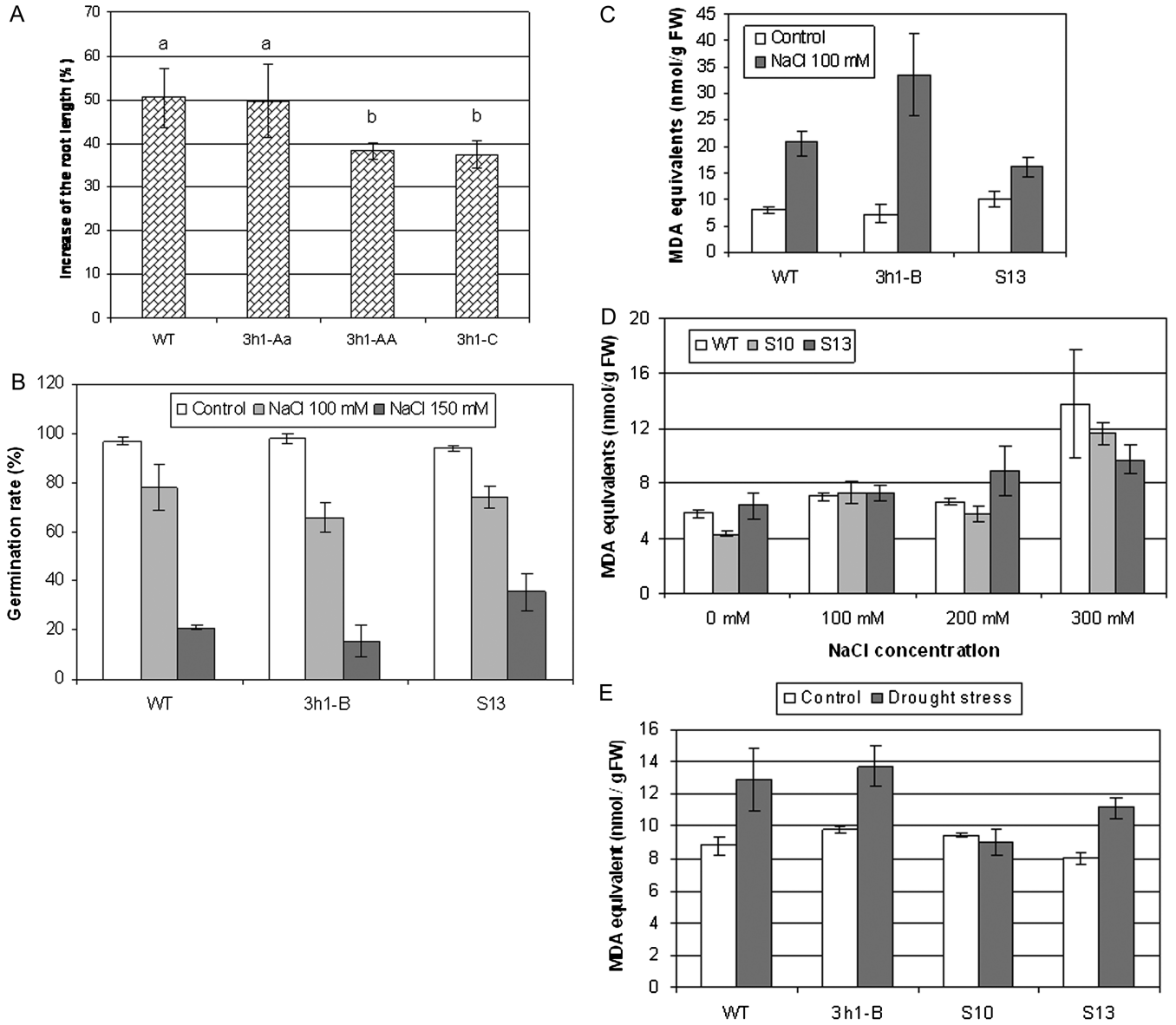

Fig. 9. Root lengths, germination rates, and MDA levels of the *ALDH3H1* T-DNA insertion mutants and of the ALDH3H1-α overexpressors subjected to salt or dehydration stress. (A) Root growth inhibition by salt. Root growth was measured as the increase in root length on MS agar medium supplemented with 100 mM NaCl. The growth on medium not supplemented with salt was taken as the reference for calculating the percentage increase in root length. *3h1-Aa* and *3h1-AA* refer to the heterozygous and homozygous knockout T-DNA insertion mutant *3h1-A*, respectively. Columns with different letters (a, b) are significantly different (Student's *t*-test; $P < 0.05$). (B) Germination rate. The percentage of seeds that germinated on salt is shown for each genotype. (C) MDA contents in 2-week-old seedlings grown on MS agar supplemented with NaCl. (D, E) MDA content in adult plants grown on soil and irrigated with different concentrations of NaCl (D) or subjected to dehydration (E). Line *3h1-B* is an ALDH3H1 knockout T-DNA insertion mutant, while S10 and S13 are ALDH3H1-α overexpressors. FW, fresh weight.

splicing transcripts and their subsequent targeting to the nonsense-mediated mRNA decay pathway is a widespread mechanism to fine tune the abundance of functional transcripts (Maquat, 2004; Lareau *et al.*, 2007; Barbazuk *et al.*, 2008; Saltzman *et al.*, 2008; Palusa and Reddy, 2010).

The data obtained for *ALDH3H1* point to the existence of an alternative promoter within intron 1 that directs expression of the alternative transcript variants. The presence of alternative promoters or enhancers within introns has been reported for many plant species (Kim *et al.*, 2006; Morello and Breviario, 2008). Several examples of introns directing the expression of a promoterless gene have been shown in rice and *Arabidopsis* (Morello *et al.*, 2002, 2006; Rose, 2008; Fojtová *et al.*, 2011). The observation that a fragment of *ALDH3H1* intron 1 is capable of directing the expression of the *GUS* reporter gene is consistent with the conclusions above. Our data indicate that intron 1 of *ALDH3H1* interferes with expression of the gene.

The alternative promoter may also direct expression of the alternative transcripts, because the alternative isoforms

are present not only in the WT but also in the T-DNA mutant *3h1-A* (Fig. 1B). Otherwise, all alternative isoforms should have been absent in this mutant line. In agreement with this, no transcript was detected in line *3h1-B*, which has a T-DNA insertion within intron 1, suggesting that the integrity of intron 1 is required for generation of the alternative isoforms. The finding that intron 1 contains an alternative promoter also supports the existence of the 5′ EST clone DR182505 of *ALDH3H1* (Alexandrov *et al.*, 2006). Our observations point to a coupling of the alternative promoter with alternative splicing events in fine tuning *ALDH3H1* gene expression. The necessity of such an alternative promoter and the function of the alternative transcripts are not understood. Chen *et al.* (2007) observed that genes involved in enzymatic reactions and cellular processes were enriched in alternative first exon transcripts in rice and *Arabidopsis*, indicating that the transcriptional regulation mediated by alternative first exons may be important for adapting to environmental changes.

The expression pattern of *ALDH3H1-α* differs substantially from that of *ALDH3H1-β*, *-γ*, *-δ*, *-ε*, and *-ζ* in roots of adult plants, which indicates that the two promoter regions are differentially activated. Alternative promoters are often responsible for tissue- or developmental stage-specific gene expression (Morello *et al.*, 2002; Kim *et al.*, 2006). Likewise, our RT-PCR analyses revealed that stress-mediated induction of *ALDH3H1-α* was restricted to leaves of adult plants, whereas induction of the alternative isoforms *ALDH3H1-β*, *-γ*, *-δ*, *-ε*, and *-ζ* occurred in both roots and leaves. The diversification of transcription start sites and/or transcripts evolved from alternative promoters or alternative splicing contributes to variations of gene expression patterns (Wang *et al.*, 2002; Tanaka *et al.*, 2009).

The comparative analysis between the *3h1-A* mutants, the WT and the *3h1-C* mutants indicated that *3h1-A* plants were affected similarly to *3h1-C* mutants in their root growth in response to NaCl. The overexpression of ALDH3H1-α diminished the accumulation of toxic aldehydes, measured as MDA in plants under stress, but the transgenic plants were as sensitive as WT plants. Compared with the results obtained from transgenic plants overexpressing ALDH3I1 (Sunkar *et al.*, 2003; Kotchoni *et al.*, 2006), this indicates that ALDH3H1 protein has a moderate role in response to drought or salt stress and may have a different function.

A general observation was the overall low expression of the alternative *ALDH3H1* isoforms in WT plants. This can be attributed to the strength of the promoters. Alternative promoters that were 100-fold different in their strength have been documented for mammalian α-amylase genes (Schibler *et al.*, 1983). The dominance of the *ALDH3H1* promoter upstream of exon 1 is probably suppressed in the T-DNA insertion mutant *3h1-A*. Insertion of the T-DNA possibly affected the integrity of some *cis* elements in the upstream promoter.

Alternative promoters and gene splicing provide additional flexibility in the control of gene regulation. It is interesting to understand why evolution has favoured the occurrence of both an alternative promoter and gene splicing in the *ALDH3H1* locus and why the activity of the alternative promoter is maintained at a low level in the WT. Generation of the alternative promoter may be an isolated event that has occurred during evolution. Landry *et al.* (2003) proposed several events that can explain the origin of alternative promoters. First, they may occur through progressive mutations, which create over time new functional motifs that can recruit the transcriptional machinery and that are in a favourable genomic position to serve as a promoter. Alternatively, they can result from a recombination event duplicating a promoter region, and subsequent mutations would affect the strength and tissue specificities of the promoters. Other possibilities are the insertion of a transposable element or genomic rearrangements near the gene, thereby creating a novel promoter. Considering that the occurrence of a novel promoter would be accompanied by the generation of novel transcription start sites, downstream promoters or transcription start sites may produce a truncated protein whose function is deteriorated or lost. In contrast, the generation of upstream mutations would affect downstream transcriptional signals and gene function. Therefore, Landry *et al.* (2003) suggested that upstream mutations would have been negatively selected during evolution and such modifications would only persist if they could resist selective pressure and become beneficial or neutral. Accordingly, the alternative promoter in the *ALDH3H1* gene probably occurred after *ALDH* gene diversification in *A. thaliana*.

Acknowledgements

This work was supported by a German Research Foundation (Deutsche Forschungsgemeinschaft)/Arabidopsis Functional Genomics Network project (BA 712/3-3) and the German Academic Exchange Service (Deutscher Akademischer Austauschdienst Dienst) (fellowship to T.D.M.). The authors are grateful to C. Buchholz for growing the plants, Dr A. Ditzer for providing the antiserum for the ALDH3H1 protein and Professor Dr H. Vaucheret for helpful comments.

References

Alexandrov NN, Troukhan ME, Brover VV, Tatarinova T, Lu Y-P, Flavell RB, Feldmann KA. 2006. Features of *Arabidopsis* genes and genome discovered using full-length cDNAs. *Plant Molecular Biology* **60,** 69–85.

An Y-Q, McDowell JM, Huang S, McKinney EC, Chambliss S, Meagher RB. 1996. Strong, constitutive expression of the *Arabidopsis* ACT2/ACT8 actin subclass in vegetative tissues. *The Plant Journal* **10,** 107–121.

Bai X, Peirson BN, Dong F, Xue C, Makaroff CA. 1999. Isolation and characterization of *SYN1*, a *RAD21*-like gene essential for meiosis in Arabidopsis. *Plant Cell* **11,** 417–430.

Barbazuk WB, Fu Y, McGinnis KM. 2008. Genome-wide analyses of alternative splicing in plants: opportunities and challenges. *Genome Research* **18,** 1381–1392.

Baulcombe DC, Saunders GR, Bevan MW, Mayo MA, Harrison BD. 1986. Expression of biologically active viral satellite RNA from the nuclear genome of transformed plants. *Nature* **321,** 446–449.

Belgrader P, Cheng J, Zhou X, Stephenson LS, Maquat LE. 1994. Mammalian nonsense codons can be *cis* effectors of nuclear mRNA half-life. *Molecular and Cellular Biology* **14,** 8219–8228.

Blencowe BJ. 2006. Alternative splicing: new insights from global analyses. *Cell* **126,** 37–47.

Campbell MA, Haas BJ, Hamilton JP, Mount SM, Buell CR. 2006. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. *BMC Genomics* **7,** 327.

Chen W-H, Lv G, Lv C, Zeng C, Hu S. 2007. Systematic analysis of alternative first exons in plant genomes. *BMC Plant Biology* **7,** 55.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for *Agrobacterium* mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16,** 735–743.

Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, Wong W-K, Mockler TC. 2010. Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Research* **20,** 45–58.

Fojtová M, Peska V, Dobsakova Z, Mozgova I, Fajkus J, Sykorova E. 2011. Molecular analysis of the T-DNA insertion mutants identified putative regulatory elements in the *AtTERT* gene. *Journal of Experimental Botany* **62,** 5531–5545.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27,** 297–300.

Jefferson RA, Kavanagh TA, Bevan MW. 1987. GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO Journal* **6,** 3901–3907.

Jeon JS, Lee S, An G. 2008. Intragenic control of expression of a rice MADS box gene *OsMADS1*. *Molecular Cells* **26,** 474–480.

Jung C, Seo JS, Han SW, Koo YJ, Kim CH, Song SI, Nahm BH, Choi YD, Cheong J-J. 2008. Overexpression of AtMYB44 enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. *Plant Physiology* **146,** 623–635.

Kim E, Magen A, Ast G. 2007. Different levels of alternative splicing among eukaryotes. *Nucleic Acids Research* **35,** 125–131.

Kim MJ, Kim H, Shin JS, Chung CH, Ohlrogge JB, Suh MC. 2006. Seed-specific expression of sesame microsomal oleic acid desaturase is controlled by combinatorial properties between negative *cis* -regulatory elements in the *SeFAD2* promoter and enhancers in the 5′-UTR intron. *Molecular Genetics and Genomics* **276,** 351–368.

Kimura K, Wakamatsu A, Suzuki Y, *et al*. 2006. Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Research* **16,** 55–65.

Kirch H-H, Bartels D, Wei Y, Schnable PS, Wood AJ. 2004. The ALDH gene superfamily of *Arabidopsis*. *Trends in Plant Science* **9,** 371–377.

Kirch H-H, Nair A, Bartels D. 2001. Novel ABA- and dehydration-inducible aldehyde dehydrogenase genes isolated from the resurrection plant *Craterostigma plantagineum* and *Arabidopsis thaliana*. *The Plant Journal* **28,** 555–567.

Kirch H-H, Schlingensiepen S, Kotchoni OS, Sunkar R, Bartels D. 2005. Detailed expression analysis of selected genes of the aldehyde dehydrogenase (ALDH) gene superfamily in *Arabidopsis thaliana*. *Plant Molecular Biology* **57,** 315–332.

Koo SC, Choi MS, Chun HJ, *et al*. 2009. Identification and characterization of alternative promoters of the rice MAP kinase gene *OsBWMK1*. *Molecular Cells* **27,** 467–473.

Kornblihtt AR. 2005. Promoter usage and alternative splicing. *Current Opinion in Cell Biology* **17,** 262–268.

Kotchoni S, Kuhns C, Ditzer A, Kirch H-H, Bartels D. 2006. Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant, Cell and Environment* **29,** 1033–1048.

Landry J, Mager DL, Wilhelm BT. 2003. Complex controls: the role of alternative promoters in mammalian genomes. *Trends in Genetics* **19,** 640–648.

Lareau LF, Brooks AN, Soergel DA, Meng Q, Brenner SE. 2007. The coupling of alternative splicing and nonsense-mediated mRNA decay. *Advances in Experimental Medicine and Biology* **623,** 190–211.

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S. 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Research* **30,** 325–327.

Li J, Li X, Guo L, Lu F, Feng X, He K, Wei L, Chen Z, Qu LJ, Gu H. 2006. A subgroup of *MYB* transcription factor genes undergoes highly conserved alternative splicing in *Arabidopsis* and rice. *Journal of Experimental Botany* **57,** 1263–73.

Li JF, Park E, von Arnim AG, Nebenführ A. 2009. The FAST technique: a simplified *Agrobacterium*-based transformation method for transient gene expression analysis in seedlings of Arabidopsis and other plant species. *Plant Methods* **5,** 6.

Luzi L, Confalonieri S, Di Fiore PP, Pelicci PG. 2000. Evolution of Shc functions from nematode to human. *Current Opinion in Genetics and Development* **10,** 668–674.

Maquat LE. 2004. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nature Reviews Molecular Cell Biology* **5,** 89–99.

Mastrangelo AM, Marone D, Laidò G, De Leonardis AM, De Vita P. 2011. Alternative splicing: enhancing ability to cope with stress via transcriptome plasticity. *Plant Science* **185,** 40–49.

Mathur J, Szabados L, Schaefer S, Grunenberg B, Lossow A, Jonas-Straube E, Schell J, Koncz C, Koncz-Kálmán Z. 1998. Gene identification with sequenced T-DNA tags generated by transformation of *Arabidopsis* cell suspension. *The Plant Journal* **13,** 707–716.

McElver J, Tzafrir I, Aux G, *et al*. 2001. Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* **159,** 1751–1763.

Missihoun TD, Schmitz J, Klug R, Kirch H-H, Bartels D. 2011. Betaine aldehyde dehydrogenase genes from *Arabidopsis* with different sub-cellular localization affect stress responses. *Planta* **233,** 369–382.

Morello L, Bardini M, Cricri M, Sala F, Breviario D. 2006. Functional analysis of DNA sequences controlling the expression of the rice *OsCDPK2* gene. *Planta* **223,** 479–491.

Morello L, Bardini M, Sala F, Breviario D. 2002. A long leader intron of the *Ostub16* rice β-tubulin gene is required for high-level gene expression and can autonomously promote transcription both *in vivo* and *in vitro*. *The Plant Journal* **29,** 33–44.

Morello L, Breviario D. 2008. Plant spliceosomal introns: not only cut and paste. *Current Genomics* **9,** 227–238.

Murashige T, Skoog F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum* **15,** 473–497.

Ner-Gaon H, Halachmi R, Savaldi-Goldstein S, Rubin E, Ophir R, Fluhr R. 2004. Intron retention is a major phenomenon in alternative splicing in *Arabidopsis*. *The Plant Journal* **39,** 877–885.

Palusa SG, Ali GS, Reddy ASN. 2007. Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins and its regulation by hormones and stresses. *The Plant Journal* **49,** 1091–1107.

Palusa SG, Reddy ASN. 2010. Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay. *New Phytologist* **185,** 83–89.

Parsley K, Hibberd JM. 2006. The Arabidopsis *PPDK* gene is transcribed from two promoters to produce differentially expressed transcripts responsible for cytosolic and plastidic proteins. *Plant Molecular Biology* **62,** 339–349.

Prestridge DS. 1991. SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Computer Applications in the Biosciences* **2,** 203–206.

Qi XT, Zhang YX, Chai TY. 2007. The bean *PvSR2* gene produces two transcripts by alternative promoter usage. *Biochemical and Biophysical Research Communications* **356,** 273–278.

Reddy ASN. 2007. Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annual Review of Plant Biology* **58,** 267–294.

Rìos G, Lossow A, Hertel B, *et al*. 2002. Rapid identification of *Arabidopsis* insertion mutants by non-radioactive detection of T-DNA tagged genes. *The Plant Journal* **32,** 243–253.

Rose AB. 2008. Intron-mediated regulation of gene expression. *Current Topics in Microbiology and Immunology* **326,** 277–290.

Saltzman AL, Kim YK, Pan Q, Fagnani MM, Maquat LE, Blencowe BJ. 2008. Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Molecular and Cellular Biology* **28,** 4320–4330.

Schibler U, Hagenbüchle O, Wellauer PK, Pittet AC. 1983. Two promoters of different strengths control the transcription of the mouse α-amylase gene *Amy-1*^{a} in the parotid gland and the liver. *Cell* **33,** 501–508.

Sessions A, Burke E, Presting G, *et al*. 2002. A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell* **14,** 2985–2994.

Sophos NA, Vasiliou V. 2003. Aldehyde dehydrogenase gene superfamily: the 2002 update. *Chemico-Biological Interactions* **143–144,** 5–22.

Stiti N, Missihoun TD, Kotchoni SO, Kirch H-H, Bartels D. 2011. Aldehyde dehydrogenases in *Arabidopsis thaliana*: biochemical requirements, metabolic pathways, and functional analysis. *Frontiers in Plant Science* **2,** 65.

Sunkar R, Bartels D, Kirch H-H. 2003. Overexpression of a stress-inducible aldehyde dehydrogenase gene from *Arabidopsis thaliana* in transgenic plants improves stress tolerance. *The Plant Journal* **35,** 452–465.

Tanabe N, Yoshimura K, Kimura A, Yabuta Y, Shigeoka S. 2007. Differential expression of alternatively spliced mRNAs of *Arabidopsis* SR protein homologs, atSR30 and atSR45a, in response to environmental stress. *Plant and Cell Physiology* **48,** 1036–1049.

Tanaka T, Koyanagi KO, Itoh T. 2009. Highly diversified molecular evolution of downstream transcription start sites in rice and *Arabidopsis*. *Plant Physiology* **149,** 1316–1324.

Towbin H, Staehelin T, Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences U S A* **76,** 4350–4354.

Wang BB, Brendel V. 2006. Genomewide comparative analysis of alternative splicing in plants. *Proceedings of the National Academy of Sciences U S A* **103,** 7175–7180.

Wang X, Su H, Bradley A. 2002. Molecular mechanisms governing *Pcdh-γ* gene expression: evidence for a multiple promoter and *cis*-alternative splicing model. *Genes and Development* **16,** 1890–1905.